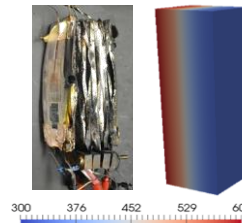
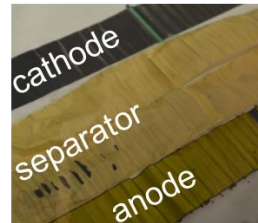
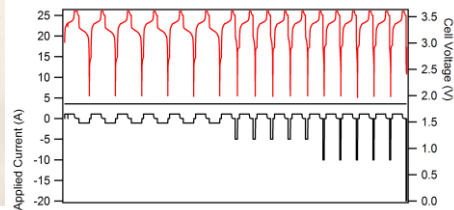


# Battery Archive – A Public Battery Data Repository



PRESENTED BY

Yuliya Preger, Valerio De Angelis (Sandia National Labs)

Presentation #603  
DOE OE Energy Storage Peer Review  
October 28, 2021

SAND2021-12729 C

## 2 PROJECT OVERVIEW



SUMMARY: Develop a user-friendly platform to make battery performance data and analysis from a variety of institutions accessible to the broader public

### SIGNIFICANCE:

- In an increasingly electrified world, many people need battery data for projects and decision making, but few have access to it
- A broad range of battery data (more than a single institution can generate) is needed for developing generalizable performance models

### ALIGNMENT WITH CORE MISSION OF DOE OE:

- Energy storage systems contribute to resilience, reliability, and flexibility of energy infrastructure
- Partnering for dissemination of battery performance data can facilitate risk assessment and 'best practices' operation of energy storage technologies

3

# PROJECT OBJECTIVES

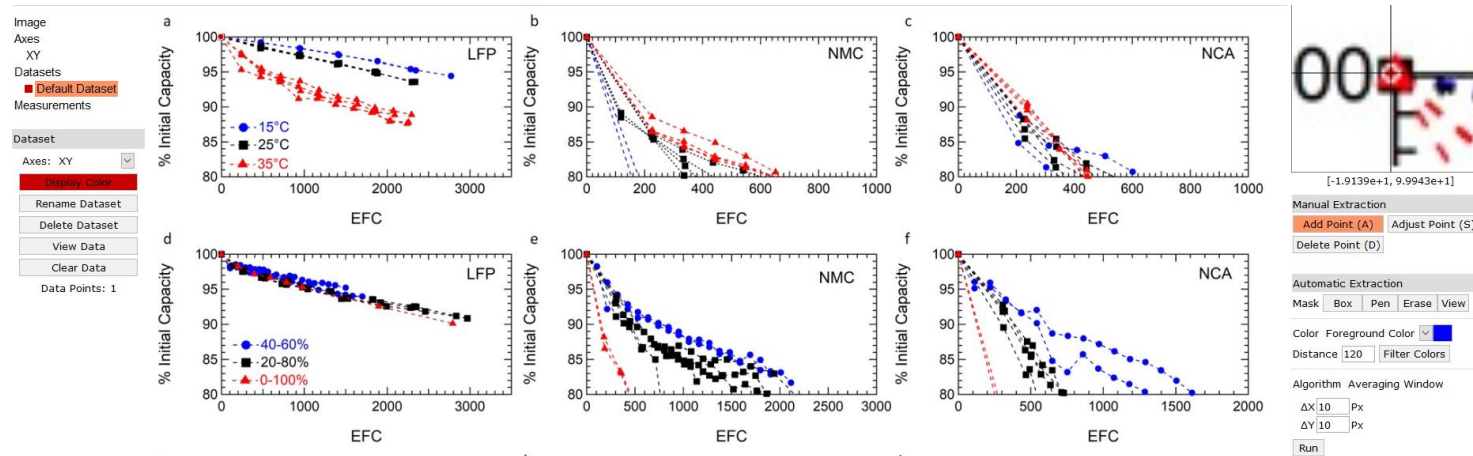


Target: simple enough for casual users to compare battery performance, but still useful for advanced modeling

User Problem	Product Feature
Access reliable battery data	Open access repository of battery data with links to studies
Determine if data is relevant before investing a lot of processing time	Immediately visualize the data
Analyze diverse battery data	Standardize data into common format and offer modeling through Jupyter Notebook
Reorganize battery data	Adopt a standard minimum set of metadata, but plan for additional optional metadata
Contribute battery data from different sources	Import data from Excel, battery testers, online repositories (Google Sheets), and databases

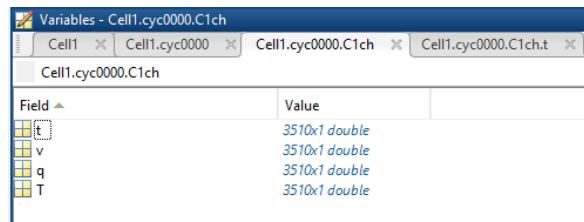
# PREVIOUS METHODS FOR SHARING BATTERY DATA

Battery data is mostly shared as figures in publications; extracting data is time consuming



Even when raw data are available, they are not standardized

## Matlab



Oxford Battery Intelligence Lab

## txt

File	Edit	Format	View	Help				
Time	Status	code	Status	category	Status	color	Pgm	code
0.000000	8	3	3	0	1	2	2	
0.940317	8	3	3	0	1	2	2	
1.954083	8	3	3	0	1	2	2	
2.950567	8	3	3	0	1	2	2	
3.945600	8	3	3	0	1	2	2	
4.940633	8	3	3	0	1	2	2	
5.937117	8	3	3	0	1	2	2	
6.952317	8	3	3	0	1	2	2	
7.948800	8	3	3	0	1	2	2	
8.942400	8	3	3	0	1	2	2	
9.938883	8	3	3	0	1	2	2	
10.952633	8	3	3	0	1	2	2	
11.944800	8	3	3	0	1	2	2	
12.952800	8	3	3	0	1	2	2	
13.940633	8	3	3	0	1	2	2	

CALCE Battery Group

## Excel

Cycle	Ind	Start_Time	End_Time	Test_Time	Min_Curr	Max_Curr	Min_Volt
1		02:10.0	22:45.3	15645.31	-0.55	0.55	1.998
2		22:45.3	40:11.8	31091.75	-0.55	0.55	1.995
3		40:11.8	57:06.2	46506.17	-0.55	0.55	1.997
4		57:06.2	26:31.6	69871.59	-1.1	0.55	1.995
5		28:31.6	34:59.9	81179.92	-1.1	0.55	1.993
6		37:00.0	43:14.7	92474.68	-1.1	0.55	1.995
7		45:14.7	51:21.0	103761	-1.1	0.55	1.998
8		53:21.1	59:21.1	115041.1	-1.1	0.55	1.995
9		01:21.2	07:15.5	126315.5	-1.1	0.55	1.995
10		09:15.6	15:06.6	137586.6	-1.1	0.55	1.996
11		17:06.6	22:52.5	148852.5	-1.1	0.55	1.998
12		24:52.5	30:37.9	160117.9	-1.1	0.55	1.994

Sandia National Labs

This makes it difficult to compare results from different studies and do larger-scale analyses



Existing resources are useful, but address different problems

## Data Management

[BEEP \(Battery Evaluation and Early Prediction\)](#): Open-source, Python-based package parses and featurizes battery cycling data to enable cycle life prediction

[cellpy](#): Open-source, Python-based package parses and enables manipulation of Arbin cycling data

[Universal Battery Database](#): Open-source, Python-based package can be used for managing Li-ion cell data.

## Electrochemical Impedance Spectroscopy (EIS)

[EIS: Measurement Model Program](#): software identifies the error structure of EIS measurements and fits custom models to the data

[impedance.py](#): Open-source, Python-based package for EIS data contains modules for data preprocessing, validation, model fitting, and visualization

## General Modeling

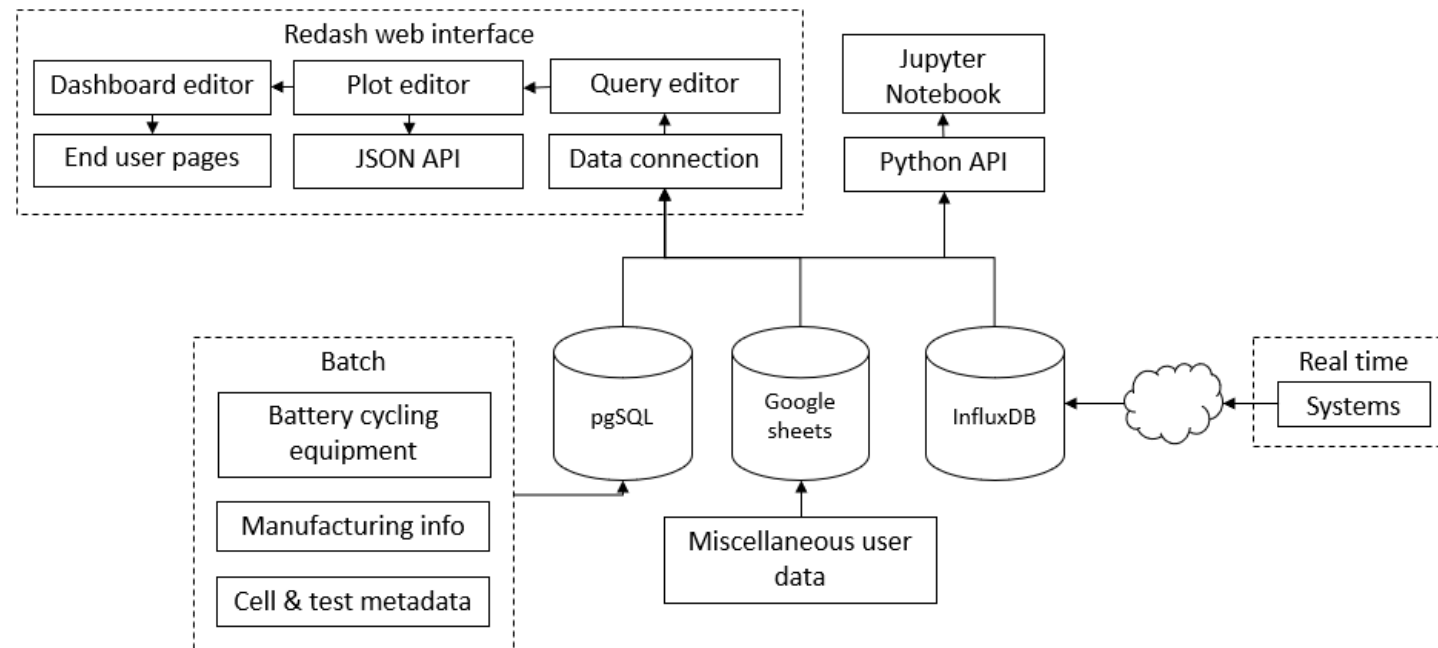
[CAEBAT](#): Various software tools for multi-scale battery modeling and pack design

[PyBaMM \(Python Battery Mathematics Modeling\)](#): Open-source, Python-based package uses various physics-based models to simulate physical properties such as voltage, concentration, and temperature of a Li-ion battery operated under different experimental protocols

# PROJECT METHODOLOGY



- We adapted well-established open-source software tools that manage/visualize data for use in the battery community
- We developed the Battery Lifecycle Framework (BLC) - a platform that provides tools to visualize, analyze, and share battery data through the technology development cycle
- BLC has four components: (1) data importers, (2) one or more databases, (3) a front-end for querying the data and creating visualizations, (4) an application programming interface to process the data



# PROJECT METHODOLOGY – LEVERAGING REDASH



Redash is an open-source extract-transform-load tool with a robust web management interface that can be used to manage data connections and generate visualizations. It offered the foundation around which we built battery-specific software tools.

## 1) Create new data source

The screenshot shows the Redash 'Type Selection' step for creating a new PostgreSQL data source. The progress bar indicates the current step. The 'Name' field is filled with 'My PostgreSQL'. The 'Host' field contains '127.0.0.1' and the 'Port' field contains '5432'.

## 2) Build data queries

```

1  SELECT
2    key || ': ' || r.cell_id as series,
3    r.cycle_index,
4    r.test_time,
5    value
6  FROM (SELECT timeseries_data.cell_id, cycle_index, test_time, json_build_object('V', v, 'C', i) AS line
7  FROM timeseries_data TABLESAMPLE BERNOULLI ({{sample_vc}})
8  where cell_id IN ({{cell_id}}) and
9         cycle_index >= {{min}} and
10        cycle_index <= {{max}}) as r
11  JOIN LATERAL json_each_text(r.line) ON (key ~ '[V,C]')
12  order by r.cell_id, r.test_time, key
13

```

{{ }}

≡

⚡

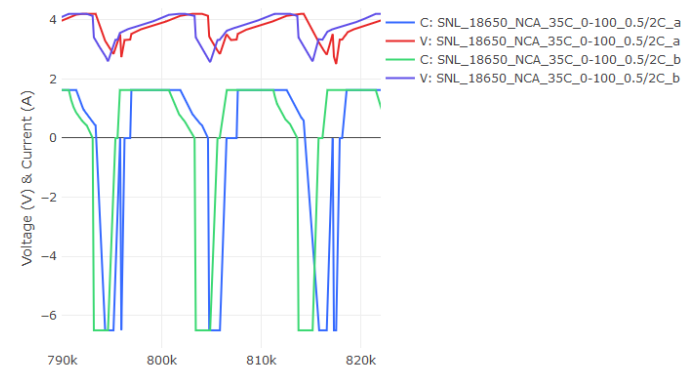
Save\*

Execute

## 3) Build visualizations

The screenshot shows the Redash 'Visualization Editor' for a 'Chart' type. The 'Visualization Name' field is filled with 'Chart'. The 'Chart Type' dropdown is set to 'Line'. The 'General' tab is selected, showing options for X Axis, Y Axis, Series, Colors, and Data Labels.

## 4) Manipulate plots



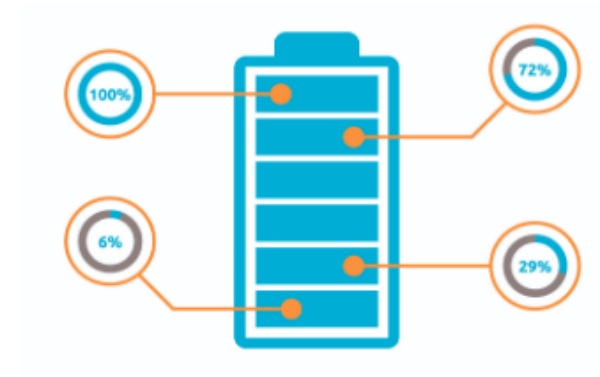
# PROJECT RESULTS: FIRST MULTI-INSTITUTION BATTERY CYCLING DATABASE



## BatteryArchive.org

A repository for easy visualization, analysis, and comparison of battery data across institutions

[View Data](#)



## Features

1

Filter battery data

Cell ID	Capacity	Temperature	Cycling Rate	Discharge Rate
UPL-000000001	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000002	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000003	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000004	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000005	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000006	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000007	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000008	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000009	2500 mAh	25.00 °C	0.5C	0.5C
UPL-000000010	2500 mAh	25.00 °C	0.5C	0.5C

Query and filter for specific experimental conditions.

2

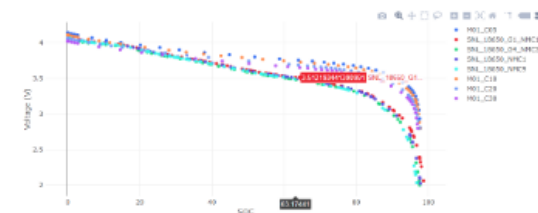
Visualize and compare data



Display battery data, including voltage curves and capacity fade.

3

Compare data with models



Apply performance and degradation models to battery data.



# PROJECT RESULTS: IMMEDIATE PRESENTATION OF COMMON BATTERY DATA VISUALIZATIONS



Search database by metadata related to cell + cycling conditions

Capacity and energy decay, efficiencies, and voltage curves automatically plotted for selected cells

BatteryArchive.org Home Cell List Studies Metadata News FAQ Resources

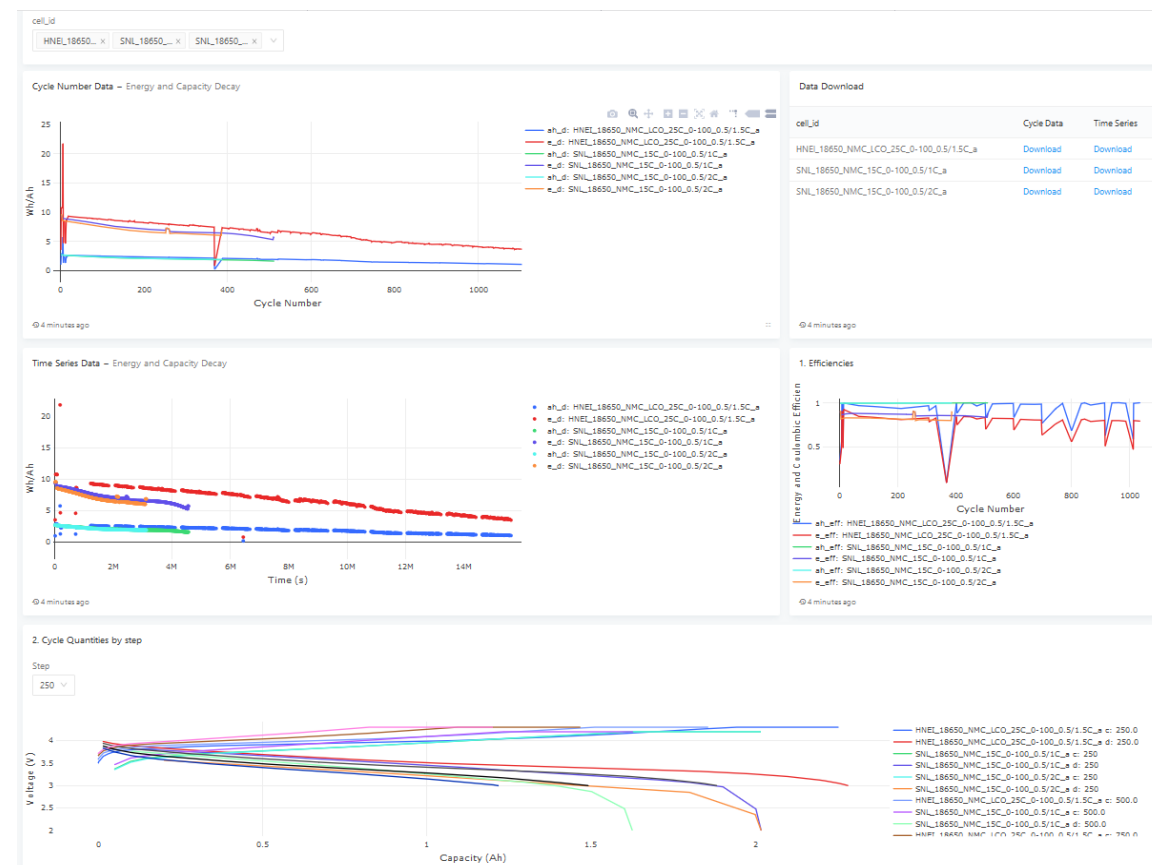
### Cell List

Cathode: LFP X NCA X NMC X +2 more  
 Capacity (Ah): 3.2 X 1.1 X 3 X +3 more  
 Temperature (C): 15 X 25 X 35 X +1 more  
 Max SOC: 100 X 60 X 80 X  
 Min SOC: 0 X 20 X 40 X  
 Discharge C Rate: 0.5 X 1 X 2 X +3 more

Source: HNEI X Oxford X calce X +1 more  
 Type: prismatic X 18650 X pouch X

Li-ion cell list

Cell ID	Cycles	Cathode	Anode	Capacity (Ah)	Type	Temp (C)	DOD	Max SOC	Min SOC	Ch C Rate	Dis C Rate	Source
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_a	1,113	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_b	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_c	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_d	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_e	1,134	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_f	1,103	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_g	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_h	1,105	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_i	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_j	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_k	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_l	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_m	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_n	1,108	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI
HNEI_18650_NMC_LCO_25C_0-100_0.5/1.5C_o	1,114	NMC-LCO	graphite	2.80	18650	25.00	100.00	100.00	0.00	0.50	1.50	HNEI



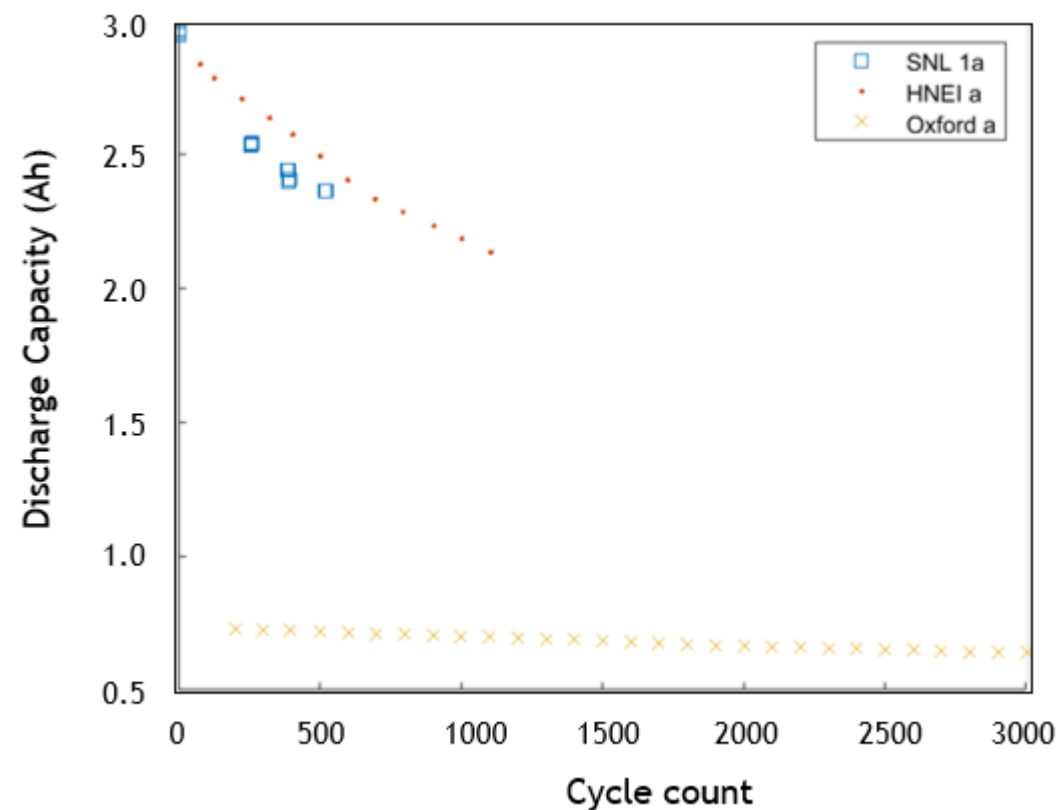
# PROJECT RESULTS: EASY TO RUN ANALYSIS ON MULTIPLE DATASETS



Development of importers to transform data into a common format enables comparison across studies and saves users many hours on pre-processing. We are pre-populating the site with generalizable analysis scripts.

Example: extracting 'capacity check' cycles from 'bulk cycling' conditions

- One cell from each of three datasets from three different institutions, each with a different original file format
- All three datasets had a different routine for capacity checks
- A single script was developed to extract the capacity check cycles from all datasets



# PROJECT IMPACT: NETWORK OF CONTRIBUTORS



## Sources of datasets & collaborators

- Datasets online
- Datasets in pipeline
- Software interoperability



# PROJECT IMPACT: USER STORIES



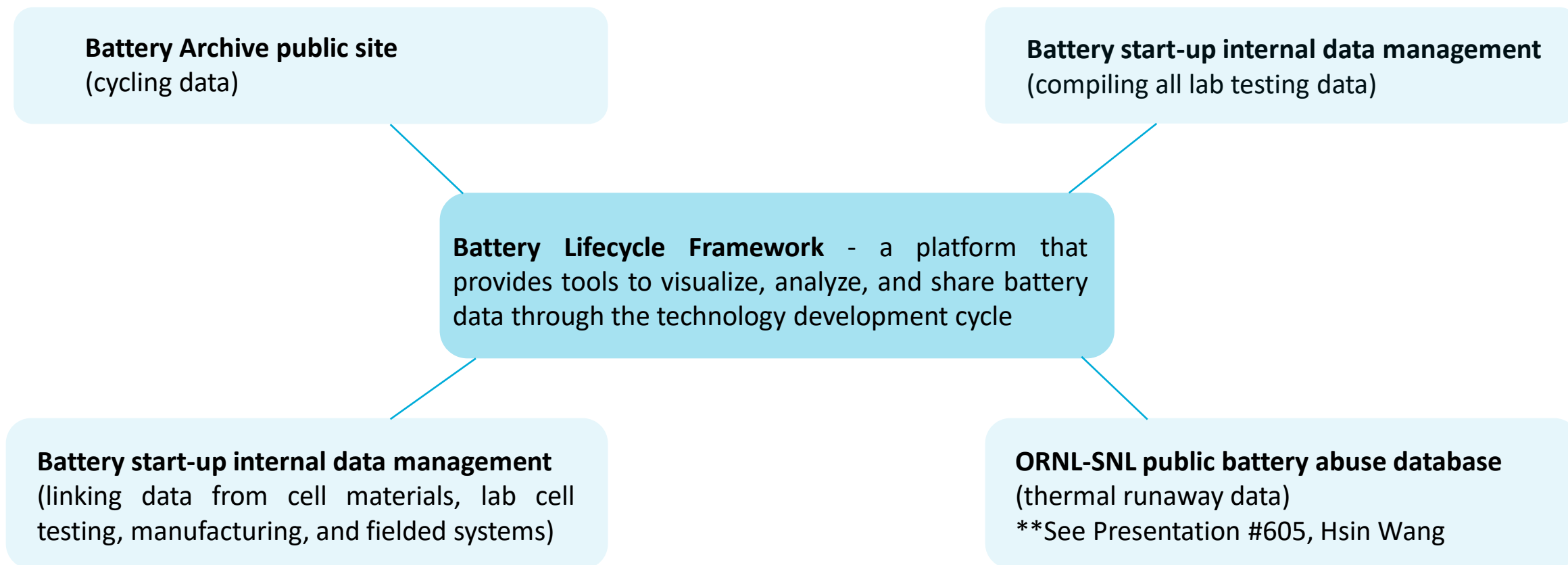
Over 5500 site users, many return visits, from over 50 countries, academia and industry

- Developers of non-battery energy storage technologies wanting to understand how their products compare to batteries under different conditions
- Representatives of utilities installing energy storage systems who are trying to get a better sense of what conditions exacerbate battery degradation
- Researchers in universities and companies who are trying to validate their battery degradation models with more data + class projects for undergrads
- Battery software start-ups which need data to test their ideas and product (new data collection would be expensive)
- Individuals at companies that already have a lot of battery data (running quick tests to try out a new idea, accessing data for different batteries, etc.)

# PROJECT RESULTS: EXTENDING THE SOFTWARE TO OTHER APPLICATIONS



The flexible software package can be extended to many other public and private battery data applications. Software compatibility allows data exchange between public and private tools as needed (e.g. some private data is shared publicly and some public data is downloaded for analysis in a private repository).



# SUMMARY OF INNOVATIONS



- First public multi-institution battery cycling database has given data access to people all over the world in a variety of industries
- Data standardization has saved users many hours on pre-processing and enables the development of generalizable battery models
- Battery Lifecycle Framework offers a complete battery-oriented open-source software package (import, visualization, analysis) that can be easily customized for different applications

# FUTURE DEVELOPMENT



- Improve the UX in collaboration with a web firm
- Enable interoperability with other open-source battery software packages (e.g. BEEP, PyBaMM)
- Develop more pre-built queries/plots for other kinds of battery data analysis
- Publish a JSON API to exchange data
- Continue uploading data to the Battery Archive public site

\*All features added to Battery Archive will be incorporated into the open-source software package

# PROJECT ACKNOWLEDGMENTS



Funded by the U.S. Department of Energy, Office of Electricity, Energy Storage program. Dr. Imre Gyuk, Program Director.

Software development: Ammar Safdari, Stephen Bauer, Irving Derin

Early framework testers at CUNY Energy Institute: Sanjoy Banerjee, Jinchao Huang, Andreas Savva, Gautam Yadav

Front end: Sam Roberts-Baca (SNL)

Database: Mark Spoonamore (SNL)

Early site feedback: Matthieu Dubarry (Hawaii Natural Energy Institute)

BEEP integration: Patrick Herring (Toyota Research Institute)

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

Public site: [www.BatteryArchive.org](http://www.BatteryArchive.org)

Battery Lifecycle Framework Preprint: <https://ecsarxiv.org/h7c24>

Github: <https://github.com/battery-lcf/battery-archive-sandbox>

For questions about this presentation: [ypreger@sandia.gov](mailto:ypreger@sandia.gov)